6
Improving the Impact of Randomized Clinical Trials

6
Improving the Impact of Randomized Clinical Trials

Throughout the course of this background paper, opportunities have been identified to improve the impact of randomized clinical trials (RCTS) on medical practice and for the expanded use of RCTS in policymaking. Potential improvements fall in the following categories: 1) the quality of RCTS that are carried out, 2) the dissemination of information from RCTS, 3) the overall system of assessing medical technologies, 4) the use of RCTS for policy decisions, and 5) the use of RCTS in specific medical fields, The following suggestions have appeared in the published literature or arose in discussions with individuals durting the course of preparing this paper.

IMPROVING THE QUALITY OF RCTS

If RCTS are to have more influence on health policymaking and medical practice, the way they are conducted needs to be improved in several ways: they should adhere to known principles of design, including statistical and other methods; they should be further improved through greater support for research in RCT methods; journal editors should impose stricter standards for RCT reports; and they should increasingly take the form of multicenter RCTS.

The Broader Application of Good Experimental Methods

Basic principles on which good RCTS depend are known. They are not always applied, however. To the extent that lack of application is a consequence of lack of knowledge of good methodology, improvements can be made at various points in the medical education system: in medical school education; in residency programs; and in continuing medical education. Outside of medical education, funding agencies, notably the National Institutes of Health (NIH), could be more assiduous in requiring good study design for funding approval, and even in providing assistance to improve deficient study designs that are submitted.

There has been some movement toward teaching quantitative methods in medical schools, but progress is slow. A suggestion for speeding up the process is to involve the American Association of **Medical Colleges (AAMC) in developing curricula for teaching research methods, including RCTS.**

The requirement for new drug approval gives the Food and Drug Administration (FDA) considerable potential leverage over the conduct of RCTS. This leverage could be used to improve the adequacy of RCTS on medical devices as well as drugs. FDA has developed, in addition to regulations, a series of guidelines for the conduct of RCTS. Adherence to these guidelines implies that results of the study will be considered as part of a New Drug Application. FDA's guidelines are quite general and set only minimal methodological standards. The guidelines could be strengthened to include standards for designing, implementing, and reporting trials. Standards for sample size, length of followup, and completeness of followup, might be considered as well as reporting requirements. Drug companies and medical device manufacturers and the groups with whom they contract to conduct RCTS are likely to be very responsive to FDA guidelines (189).

In part, a lack of faculty qualified in quantitative methods may hamper the teaching of these methods in medical schools. NIH has *a* program of career development awards in medicine, but none in the field of biostatistical methods. Making such awards might further the teaching of quantitative methods in medical schools (255).

Providing assistance for designing sound RCTS by granting agencies is not a new idea. The National Eye Institute (NEI), in the early years of encouraging RCTS, made small planning awards to those with good ideas, but in need of statistical and methodological assistance for RCTS. Such a program could be targeted to areas of medicine in which RCTS still are not widely used.

Improving Statistical Methods Through Research

The application of known statistical principles in trials would go a long way toward improving them. There is also scope for improving the methods themselves. The RCT is a relatively recently developed method, and deserves to be developed to the fullest.

The Federal research establishment does not now systematically support research to develop biostatistical methods. NIH has no study section to review grant applications in biostatistics and clinical trial methodology, and therefore relies on ad hoc groups. **As** a result, these groups may not be made up of those most qualified to review the grants received. **A** permanent study section would likely be more carefully chosen, and its existence might encourage more grant proposals to develop innovative methods in clinical research. Further, improving RCTS will depend on advances in biostatistical methods.

Applying Stricter Editorial Standards

Because publication is a critical part of the RCT process, and publications are important to the careers of researchers, journal editors wield a powerful tool in their standards for acceptance. Many have argued that these standards should be more rigorous. Curtis Meinert, the editor of Controllecl ***Clinical Trials,*** proposes that the following information should be required in a report for publication (159):

- the source of funding for the trial and an indication of whether the reported results are a subgroup of a larger data set;
- a list of the treatment groups and the rationale for the choice of treatments;
- a description of the method to allocate patients to treatment groups, including reference to the blinding used in each group (i.e., none, single or double blinded);
- the safeguards used in the trial to protect patients informed consent and privacy;
- the criteria used to exclude patients from the trial;
- the criteria used to include patients in the trial;
- the rationale for the number of patients studied, including a statement of assumptions used in calculating the sample size;
- a statement of the length of time required to complete patient enrollment;
- a description of the population from which patients were selected;
- a description of the baseline and followup examination schedule;
- a specification of the key outcome variable(s);
- the descriptive information on the baseline comparability of the treatment groups;
- the number of patients assigned to each treatment group;
- the level of patient compliance achieved in each treatment group;
- the number of patients followed to the end of the study or to death;
- the number of deceased patients;
- the number of patients unable or unwilling to return for followup examinations, including a count of the number who could not be located at the end of the study;
- a description of quality control procedures used in collecting data;
- a description of the methods of analysis, including an indication whether the reported p values resulted from a single or repeated evaluation of the data; and
- a discussion of the power of the study.

Encouraging Multicenter RCTS

Multicenter RCTS should be encouraged in situations where increased sample size and a more heterogeneous population are assets. Strategies to overcome some of the difficulties of multicenter trials should be developed.

Carrying out multicenter trials requires that a large number of investigators cooperate, however, and the present incentives for individuals to do so are low, regardless of their interest in the study. Reports of multicenter RCTS often cite the author as the cooperative group or may list a dozen names, sometimes at the report's end. Such forms of citation do little for the professional standing

of researchers in academic settings. Meinert (159) suggests the following to overcome some of these disincentives:

1. Encourage investigators to participate by recognizing participation in promotion criteria for academic faculty.
2. Allow greater flexibility for participating investigators to carry out related investigations which they can publish under their own names.
3. Award greater recognition to the field of clinical trials as a professional activity and not just as an adjunct to treating patients.

Noting the contributions of community hospital physicians in recent trials (ch. 4, "Multicenter Trials," and ch. 5, "Impact of the Cooperative Oncology Groups"), Cease (38) argues that such participation is, in fact, continuing medical education (CME). He suggests that CME credit be awarded for a certain level of participation, to serve both as recognition of achievement and as an incentive to participate.

IMPROVING THE DISSEMINATION AND USE OF RCT RESULTS

The results of RCTS can be useful in several ways. One well-designed, well-conducted RCT can provide convincing evidence for a change in practice. In that case, the results should be known to clinicians so that they may change their behavior accordingly. The results of another RCT might not be so unequivocal. They might not be the basis for altering practice immediately. If there are enough other trials on the same subject, the results taken together might suggest a clearer answer. That situation calls for some type of synthesis, perhaps a meta-analysis of RCTS. Publication of the synthesis results might then be the basis for changing clinical practice.

An RCT may confirm that current practice is indeed effective, or more effective than a newer practice, and those results should be known to physicians in the appropriate fields.

In addition to providing guidance for medical practice, RCTS may contribute to further research, either in the design of future RCTS or in other types of research. In that case, it is researchers who will benefit from knowing the results of the RCT.

Finally, information about patient treatment techniques, other than the final result, is generated in RCTS.

OptimaI strategies for disseminating information from RCTS will differ depending on which group needs to know about the results, and what aspects of the results are most relevant. Two basic approaches are needed:

1. an active dissemination effort, trying to reach those who need to know with the results, and
2. facilitating access to RCT results for those who want to find out.

The traditional and still most important method of disseminating research results of any kind, including those of RCTS, is through publication in technical journals. This may be sufficient for trials that are not of great clinical significance. For those which clearly point the way for changing medical practice, however, a single publication, even in the most prestigious medical journals, may not reach those who need to know, namely the practitioners in the field of the trial or general practitioners who sometimes or frequently work in the areas. In some cases, interesting results in treating diseases of high public visibility may lead to publicity in the mass media, but such occurrences are rather rare. Medical news publications report on a greater proportion of research results of clinical interest. Beyond those routes, there must be greater initiative on the part of investigators and perhaps funding agencies to disseminate findings from RCTS.

Pharmaceutical companies make the most direct use of RCT results in advertising their products. Implicit in their statements about safety and efficacy is the backing of RCT results. They advertise both in widely read subscription journals and in widely distributed "throwaway" publications. In addition, their representatives personally visit physicians and institutions. Together these public

relations achieve widespread awareness of a company's products.

FDA might also draw clinicians' attention to RCT results if they more formally included RCT results in the Physicians' Desk Reference (PDR) drug inserts. Inclusion of a brief account of supporting RCTS, indicating the methods, results, and limitations of the trials would provide clinicians with a basis for their own critical analysis before prescribing a drug (189).

Government and private funding agencies probably cannot match the efforts of pharmaceutical companies, and to do so might not be desirable. Nevertheless, they could greatly improve in this regard. The National Heart, Lung, and Blood Institute (NHLBI) leads such funding bodies in disseminating results (ch. 5), Its use of the medical news media, workshops, and meetings could serve as a model for other organizations.

The medical specialty societies also help disseminate information. Most active at present is the American College of Physicians (the association of physicians who have demonstrated competence in internal medicine). These societies should be encouraged to educate members both about RCT methods and about the results of specific RCTS.

The institutes of NIH, to varying degrees, also disseminate information by holding meetings at **the** NIH **campus, sponsoring** sessions at meetings of specialty societies, and sponsoring and disseminating the results of consensus development conferences.

The National Cancer Institute (NCI) has begun a program to facilitate access to active trials in clinical cancer research. The "PDQ" system is an international computerized data base accessible to patients and physicians, containing protocols of clinical research (see ch. 5).

Chalmers and his colleagues have begun a major effort to facilitate access to RCT results in all fields of medicine. Having collected published reports of RCTS for a number of years, as of 1982 a total of nearly 3,000, they have begun computerizing this information so that investigators and clinical physicians can have ready access to data on RCTS in specific areas. This is not possible through any existing data base. Included in each entry is an evaluation of the trial by Chalmers' quality index (ch. 4). The system will facilitate the synthesis of results from trials in many fields.

With the proliferation of personal computers, data bases such as Chalmers has established and NCI'S PDQ system should be available to practicing physicians. Funding agencies and the preparers of data bases could profitably undertake efforts to ensure that clinically relevant research, including RCTS is readily accessible to clinicians with personal computers.

Probstfield and his colleagues (185) have identified a failing in dissemination of information from RCTS which has rarely been addressed. It is that "the methodological knowledge gained from clinical trials cannot at present be systematically transferred to clinical practice. " The areas that Probstfield and his colleagues have identified in which clinical trial methods can contribute to clinical medicine are: clinic operations and management, the quality control of clinical practice, patient adherence to therapeutic regimens, and staff education. Information about these subjects may be available even before the trial is over. The authors suggest some steps that would improve the access to and use of information from clinical trials:

- **a** computerized retrieval system at some central source for clinical trials methods must be developed, maintained and consistently updated with appropriate cataloging of new developments;
- scientists in clinical trials must make additional efforts to recognize and to highlight in specific publications the methodology which is relevant for clinical practitioners;
- a systematic transfer of the clinical trials methodology literature to that literature read by the clinical practitioner is crucial. This transfer may require brief summaries of methods published regularly in journals with appropriate circulation and readership; and
- facilities on a national or regional basis must be developed to train clinical practitioners in methods validated in clinical trials.

IMPROVING THE ASSESSMENT OF MEDICAL TECHNOLOGIES

The results of RCTS should have the greatest impact possible. This entails developing a rational means to set priorities in funding research given the limited dollars available. The priority criteria should take into account which technologies are important for health policymaking and medical practice.

NHLBI's decisionmaking procedure for large-scale RCTS is one model for a mechanism to set priorities (see ch. 5, "NHLBI and RCTS"). Bunker and Fowles' (27) "Institute for Health Care Evaluation" (IHCE) proposes another model for this mechanism to improve the evaluation of medical technologies in all its phases (see ch. 3). one important function of IHCE would be to set research priorities.

Perry (178) proposes that a "Center for Assessment of Health Care Technology" be established in the private sector. Like IHCE, this Center would be a nonprofit organization funded by several sources: "private foundations, private third-party payers and health insurance alliances, group health and hospital associations, and corporations and labor unions with major health insurance programs for employ ees. " Perry adds, "it is also conceivable that funds could be obtained under contract from HCFA [the Health Care Financing Administration] for evaluations to be used in coverage decisions and from other Federal or State agencies requiring similar services. " Though Perry applauds related activities in the private sector such as the Clinical Efficacy Assessment Project of the American College of Physicians (ch. 3) and other efforts sponsored by the medical community, he thinks they cannot replace the impartial assessment that is possible by an organization without special interests—e.g., the proposed center.

Suggestions have been made to increase the efficiency of the process leading up to clinical trials. This would require the earlier identification of technologies that will need assessment and the improved use of information gathered prior to any RCT. If a new procedure is first tried on patients at various locations around the country, for instance, the data collected on each case could profitably be standardized and pooled, and perhaps placed in a data bank. None of these procedures are generally followed today, and many more patients than those required may undergo the procedure before one center or group has sufficient data to plan a good trial.

Mosteller and Weinstein (164) have proposed a method to evaluate the costs, risks, and benefits of clinical trials before they are carried out. Their technique is proposed to improve the rationality of spending for medical research and evaluation. In essence, the evaluation attempts to predict what the impact of doing a trial may be and with that information to decide whether the trial would be worthwhile. The authors lay out a large number of assumptions and uncertainties in formulating their model. One of its valuable aspects is that it forces a wide range of probable impacts to be considered, not only the potential benefits and risks of the procedure, but also the potential value of new knowledge gained about the disease, clinical trial methods, and health *services* delivery, for example. Such issues as possible misapplication of the procedure, the probability of widespread diffusion of a technology before the study is completed, and other relative unknowns figure in the evaluation.

An additional benefit of the evaluation is that it facilitates actual assessment of impact after a trial is finished, a task which has seldom if ever been accomplished with total success.

THE USE OF RCTS IN POLICY DECISIONS

Some have suggested that the trend of using RCT results in making policy should be encouraged. In large part they refer to decisions about coverage of medical services by third-party payers, both public and private. RCT results might be more useful for policy decisions if there were

greater interaction between third-party payers and funding agencies. This would help to focus RCTS on health issues directly relevant for policy, and more generally, to make all RCTS more relevant to policy. The latter could be accomplished by including components on cost, for instance. Contributions to funding RCTS, discussed in the section below, might help in this effort.

At lower levels of policymaking, RCT results could be used more extensively by hospitals and other medical institutions in decisionmaking about their services.

Funding of RCTS

NIH spends more money funding clinical trials than any other institution in the United States, and perhaps, in the world. In the last year for which figures are available, NIH spent 4.3 percent of its total budget on clinical trials (not all are RCTS; see ch. 2). In 1975, it spent 5 percent of the total budget on clinical trials. The trend since 1979 is unknown, though there is reason to believe the share spent on clinical trials has diminished (78). Even at the 1979 level, "136 million of an approximately $3 billion total budget for NIH, shows a rather small commitment to testing the results of years of basic and applied research" (17).

Apart from increasing NIH funds for clinical trials, funding can be increased to the extent the costs of RCTS can be distributed more fully within the health care system. Third-party payers currently reimburse for some costs of patient care and hospitalization in RCTS. That share could be increased (see ch. 2 for a full discussion of RCT funding by third-party payers). Some progress has been made, and efforts are under way to facilitate greater participation in RCT funding by health insurers.

For the first time, as a result of the 1983 Social Security Act Amendments, HCFA will be allowed to fund RCTS. Presumably they will use that capability to answer questions of direct policy relevance to the Medicare and Medicaid programs. Not only does HCFA have the opportunity to provide useful information, but their activities, if successful, may stimulate similar commitments among private third-party payers.

IMPROVED USE OF RCTS IN SPECIFIC FIELDS

Suggestions have been made to extend and improve the use of RCTS in specific areas of medicine and for specific types of technologies. These are discussed below.

Surgery

The uses and limitations of RCTS in surgery are discussed in chapter 5. The recommendations made by Bunker and colleagues in Costs, Risks, and *Benefits of Surgery* (28) are reproduced here:

> **Recommendation 1**
>
> Appropriate studies of the effectiveness of surgical treatment should be carried out for selected conditions, particularly those where uncertainty leads to professional disagreement.
>
> . . . *Improving techniques for evaluation.* At the same time that studies using currently available methods must go forward, we have seen the need to improve our ability to conduct these urgently needed studies. A major problem is our presently inadequate information system. Separate records are kept for each patient by each physician or institution caring for him. In 1977 it is possible to identify outcome as related to an operation or other treatment only if the treatment and the observed outcome occur during a single continuous hospitalization. Even under these circumstances the standard medical record is not designed for easy information retrieval or the pooling of information across patients to study populations. It is frequently nearly impossible to document the treatment and health status found at previous examinations, especially if a different hospital or physician were responsible. Existing data cannot determine long-term outcomes or the end-result of surgery. Thus we are unable to find out, except for selected conditions such as malignant tumors and end-stage renal disease, how many patients survive one or more years after a particular operation. We cannot determine how many patients have been relieved of the condition leading to the operation, or how many fully

recovered from the effects of anesthesia and surgery and been able to return to full, pre-illness activity.

We are now able to perform useful cost-risk-benefit analyses, but present techniques need to be improved; for example, we are probably not sufficiently aware of second order effects or unanticipated consequences of proposed new policies. Perhaps we can learn to anticipate such "unanticipated consequences. " Careful work still remains to be done on methodology of experimental design. It is not sufficiently widely recognized how long it takes to design an informative clinical trial or how difficult it is to execute the design once it has been chosen. We do not yet know enough about randomized trials and their consequences, their weaknesses, strengths, and costs compared with their alternatives. We still are not sure enough of when we should trust an observational study. We do not know how to combine epidemiology and observational and experimental information. We have not dealt with the ethical issues surrounding human experimentation and are still shouting at one another from fixed positions. We have not reviewed the complexities of our ethical problems in enough detail or sophistication.

Recommendation 11

Our grasp of the components of cost-benefit analysis and their interrelations, the values of the various data gathering techniques, and our understanding of the ethics of data gathering must be improved by theoretical and empirical work and by continued discussions in the public forums.

. . . *Improving medical capabilities for evaluation*. In addition to assessing the efficacy of many existing treatments, we need to develop a policy for the introduction of new medical and surgical technology. Thus among the studies encouraged in Recommendation 11, we would include further historical studies of past successes and failures. We call particular attention to two recently published studies. One, the "Study on Surgical Services for the United States' '(l72), includes a survey of the major surgical advances of the past quarter-century and the research on which these advances were based. The second, entitled "Scientific Basis for the Support of Biomedical Science" (54), examines in detail the research basis for recent advances in the surgical and medical treatment of cardiovascular and pulmonary diseases. Studying only successes or failures can have weaknesses that a balanced approach may avoid.

Even when the technology and data may be available, the current methods need to be more widely understood in the medical research and medical policy communities as well as among medical students and their teachers. Naturally, we cannot expect all to be experts. But physicians themselves must be better educated in the analytic techniques necessary for them to make a more informed discrimination among therapeutic programs or techniques, and they must be educated in the economic, social, and epidemiological principles of medical care which will allow them to participate as leaders of society in advising on or helping to make priority decisions.

Recommendation III

These principles of cost-benefit evaluation should be included as an integral part of the medical school curriculum; and their application to the assessment of the efficacy of medical care should be incorporated into clinical practice and continuing medical education.

We note in particular that medical students at the beginning of their clinical training may feel little pressure to know much about the design of clinical trials or of policy analysis. Later, when working in the hospital and trying to read and appraise results presented in research papers or in participating in research, knowledge of these matters absorb the young physician's attention. Thus, we stress continuing education.

Improving public understanding. In addition to educating itself, the medical community has an obligation to inform the public. Here we would note a distinction made by the sociologist Paul Lazarsfeld between advising and deciding. After data are gathered by good methods and carefully analyzed, the scientist or physician needs to advise the client, here the community, about the findings. The community takes this advice and tempers it with political, legal, social, and moral considerations and then decides. We should improve our advice so that it will be useful in the decision process.

Recommendation IV

Information on outcomes as well as costs of medical care should be routinely formulated in a manner suitable for presentation to the public.

Cancer Research

The use of RCTS in cancer research could be improved through better statistical analysis of the potential value of a trial, and through directing them more frequently to research in cancer prevention.

Zelen, Gehan, and Glidewell (258) suggest that the following conditions be met for a trial to be done:

1. Do not initiate a definitive clinical trial unless there is a reasonable a priori probability greater than 0.05 that a clinically important gain may exist. One way of interpreting this rule or behavior is to carry out pilot studies before launching a definitive study. If the pilot studies are encouraging, then proceed with a large comparative study.
2. Comparative trials should be planned with a minimum of 100 to 200 patients per treatment. Trials with fewer patients are likely to produce more false positive results than true positive results.
3. All positive results should be independently confirmed. This will lower the false positive rate and raise the true positive rate. Physicians in practice should exercise caution in adopting a new therapy if there is no independent confirmation.

Greater emphasis on cancer prevention is warranted in RCTS. The first major trial in primary prevention is now under way. Sponsored by NCI, it is testing beta carotene, a precursor of vitamin A, as a cancer inhibitor. One important cancer screening technique, the use of mammography to detect breast cancer, has been carefully evaluated in RCTS. Several trials of lung cancer screening are now ongoing. The survival rate of those with the most common types of cancer—lung, gastrointestinal, and breast cancers—has not improved greatly since the **1950's (226).** Thus, the detection and treatment of cancer at its early stages seems a reasonable immediate goal. Though admittedly expensive and administratively complex, the larger trials necessary to evaluate screening procedures would be worthwhile. They might compare favorably in the information they produce with large-scale secondary prevention trials in cardiovascular disease.

CONCLUSION

This paper has reviewed the available literature about the impacts of RCTS. The use of RCTS themselves is a relatively recent development, beginning only in the middle of this century and still gaining in popularity. Concern about the impacts of RCTS has come even more recently, and ideas for improving or increasing these impacts have been little voiced, Based on the small literature now available, additional effort could be profitably directed toward understanding the impacts of RCTS, and devising methods for maximizing their usefulness in health policymaking and in influencing medical practice. RCTS could play a greater role in the national use of medical technology at all levels of decisionmaking.